SIMILARITY MEASURE FOR RETRIEVAL OF QUESTION ITEMS WITH MULTI-VARIABLE DATA SETS

SITI HASRINAFASYA BINTI CHE HASSAN

UNIVERSITI TEKNOLOGI MALAYSIA

ABSTRACT

In designing test question items assessment, similarity measures have a great influence in determining whether the test question items generated semantically match to the learning outcomes and the instructional objectives. It has been realized that to carry out an effective case retrieval of question items, there must be selection criteria of questions' features that considerably meet the specifications and requirements of learning outcomes as well as instructional objectives that are set by academician. In this case, each question item consists of multi-variables data type namely, Bloom level, question type, discrimination index and difficulty index. To retrieve the semantic similar question items, it strongly depends on the correct definition of the case representation as well as similarity measure. In other words, the representation of data must reflect the characteristic of data type before the appropriate adapted similarity measure approach can be applied to measure the degree of similarity values. In this case, Bloom was transformed into normalized rank data before Euclidean distance similarity measure was applied. Meanwhile, question type was converted into binary, 0 and 1 before Hamming distance was applied to calculate its similarity value. Both difficulty index and discrimination index used the concept of fuzzy similarity measure, whereby their index ranges were adjusted and expressed in trapezoidal fuzzy numbers, respectively. Lastly, these approaches were aggregated together to produce one single similarity value of question item.

ABSTRAK

Dalam menggubal soalan-soalan ujian penilaian, pengukuran kesamaan mempunyai pengaruh yang besar dalam menentukan samada soalan-soalan ujian yang telah dijana benar-benar bertepatan dengan hasil akhir pembelajaran dan objektif pengajaran. Ia diakui bahawa, untuk menjana soalan-soalan ujian yang berkesan, pemilihan soalan perlu dibuat berdasarkan kriteria-kriteria tertentu yang memenuhi spesifikasi dan keperluan hasil akhir pembelajaran dan objektif pengajaran yang telah ditentukan oleh pengajar. Dalam kes ini, setiap item soalan terdiri daripada pelbagai jenis data iaitu, *Bloom*, jenis soalan, indeks diskriminasi dan indeks kesukaran. Untuk memperolehi soalan-soalan yang benar-benar serupa dari segi semantik, ia sangat bergantung kepada ketepatan perwakilan data dan pengukuran kesamaan. Dalam erti kata yang lain, perwakilan data hendaklah menggambarkan ciri-ciri bagi jenis data tersebut sebelum pengukuran kesamaan yang sesuai digunakan untuk mengukur darjah bagi nilai-nilai kesamaan. Dalam kes ini, Bloom ditukarkan kepada *normalized rank data* sebelum pengukuran kesamaan *Euclidean distance* digunakan. Manakala, jenis soalan ditukarkan kepada sistem angka perduaan, 0 dan 1 sebelum Hamming distance digunakan untuk mengira nilai kesamaan. indeks diskriminasi dan indeks kesukaran menggunakan konsep pengukuran kesamaan kabur di mana, julat bagi indeks masing-masing diubahsuai dan diterjemahkan ke dalam nombor kabur dengan graf berbentuk trapezium. Akhirnya, kesemua pendekatan ini digabungkan bersama-sama untuk menghasilkan satu nilai kesamaan bagi satu item soalan.

TABLE OF CONTENT

LIST OF GRAPHS

CHAPTER 1

PROJECT OVERVIEW

1.1 Introduction

Nowadays, similarity measure approach is one of the most interest areas in retrieving closely similar cases that stored in database. It has been reported that many case-based application systems that deal with a great quantity of data manipulation tend to apply particular similarity measurement techniques in retrieving the similar past cases. In designing question items, it has been admitted that revising the existing similar question items considerably is more efficient than creating a new question. Thus, similarity measure has a great influence in evaluating whether the stored past questions retrieved are approximately similar to the test blueprint. In other words, once the test blueprint criteria of question have been defined, several past questions that considerably similar to the test blueprint criteria will be chosen to create a new question.

In particular, case-based application systems are designed to match cases stored in a database with new cases. In other words, this system uses past cases namely case bases as a basis for dealing with novel problems. In this situation, a case is represented as a test question together with the associated certain criteria for each question items such as Bloom’s taxonomy level, question type, difficulty index and discrimination index. Whenever academicians want to prepare a set of test question, they will compare the current question items with the similar past question items that have been stored in a database. It means that similarity measurement considerably has great influence on the retrieval cases conditions that suit to the new desired question.

In general, in order to find the test question(s) that considerably most similar, there are two main processes involved. Initially, a new question item that is desired to be created will be checked against the existing past question items stored in case base. This process can be done by evaluating the difference of distance between the desired question item and the existing question items stored. Afterwards, similarity value between the corresponding cases will be measured, producing the suggested similar case solutions. However, each question items consists of several multi-variants of data types. Therefore, an appropriate similarity measure approach that will be incorporating with multi-variables data sets need to be proposed, particularly.

1.2 Background of Problem

Retrieving of similar multi-variables question items is the problem being focused by this research. In every semester of each year, academicians need to prepare questions for various purposes of assessments in order to determine whether the students have achieved certain learning objectives. Since the main purpose of generating test questions is to determine whether the corresponding objectives have been achieved, the test question items generated should match the learning outcomes and the instructional objectives. It has been realized that the process of preparing and designing test questions based on relevant purposes of test is always time consuming, redundant and difficult to implement. Moreover, it is the fact that some of the question items that have been used for that assessment may be reused or revised for future purpose of assessments.

Usually, a typical question item generation is done through random generation. However, the randomized approach normally does not consider the learning objectives and other criteria set for the particular assessment. In fact, according to the outcome based education, each question test should have certain bloom taxonomy of cognitive objectives that indicates the level of student’s thinking and other specific criteria which describe the question items such as the question type, the difficulty index, and the discrimination index. Besides, it has been observed that a usual searching operation is only based on finding an exactly matched question. However, this way of searching is not appropriate for finding question items that are to be reused and revised. Therefore, there is a work done which focus on the similarity measurement method to retrieve similar question items from the question bank. In order to find the similar question item, most of the retrieving works implement traditional approaches such as the Hamming distance and the Euclidean distance techniques. However, these techniques only can be applied for certain feature of data type. For example, Hamming distance is suitable to be used for binary data type whereas, the numeric data type is applicable with Euclidean distance. Since there are certain data types in question items that consist of several level of categories, Fuzzy similarity measure also require to be applied in measuring the similarities among the retrieved question items. Thus, these three similarity measure approaches need to be proposed in order to measure the multi-variants of question items’ data types.

1.3 Statement of the Problem

In a conventional method, the process of generating question items from the question bank is performed through randomized approach and exact matching. Since the purpose of generating and retrieving question items are for reuse or revise, these conventional approaches are considerably not a good solution. The similarity measure is seen as a promising approach in which it involves the process of finding the most similar case to the query. However, since the question items consist of multi-variants of data types that need to be considered as well, Hamming distance, Euclidean distance and Fuzzy similarity measure are the applicable to be applied and aggregated together in retrieving similar cases of question items.

It has been reported that traditional similarity measure technique can only handle features with real-value and characteristic feature values. Unfortunately, in the real world situation, case features are often vague or uncertain. The most common example is, one of the features of cases may be described by such linguistic terms such as *low*, *medium*, and *high*. Then, for implementing the process of case matching and retrieval, one needs to define an appropriate metric of similarity. The traditional definition of similarity is obviously not valid and at least not effective to deal with this difficulty. Hence, it is a challenge to build an effective question items generation system that meets pedagogical aspect of learning. Moreover, question items should match the learning outcomes, as well as the conditions determined by the instructional objectives. Therefore, there is a need to study the feasibility of similarity measure approach for retrieving the closely similar multi-variables question items that satisfy the specifications and requirements of learning objectives.

1.4 Project Aim

The aim of this project is to investigate the feasibility of similarity measure approach for retrieving the closely similar multi-variables question items that meet the specifications and requirements of learning outcomes as well as instructional objectives that are set by academicians.

1.5 Objectives of Project

There are several objectives that would like to be achieved in this project, shown as follow:

i. To study the feasibility of case representation approaches for similarity measurement of multivariate data types.

ii. To analyze the similarity measure retrieval based on certain criteria for each question item such as the Bloom's taxonomy level, question type, difficulty index and discrimination index.

1.6 Scopes of Project

Scope can be illustrated as a project's boundary that guides the limitation of project implementation. Scopes of this project are explained as below:

i. This project focused on the three similarity measure approaches namely Euclidean distance, Hamming distance and Fuzzy similarity measure that integrated together to measure the similarity values of multi-variants question item data types.

ii. The work implementation only considered on the case representation format and similarity measure retrieval process in order to generate question items that are closely similar to the query question items.

REFERENCES

Armengol, E., Esteva, F., Godo, L., and et. al. (2004). *On Learning Similarity Relations in Fuzzy Case-Based Reasoning*. Journal of Springer-Verlag Berlin Heidelberg 2004, 14-32.

Armengol, E., and Plaza, E. (2001). *Similarity Assessment for Relational CBR.* Journal of Springer-Verlag Berlin Heidelberg 2004, 44-58.

Bourbakis, N. G. (1998). *Artificial Intelligence and Autamation.* Advanced Series on Artificial Intelligence – Volume 3, World Scientific Publishing Co. Pte. Ltd.

Butler, S. M., and McMunn, N. D. (2006). *A Teacher's Guide to Classroom Assessment. Understanding and Using Assessment to Improve Student Learning.* 1st Edition, Jossey-Bass Teacher, A Wiley Imprint.

Chen. S. J. and Chen. S. M. (2003). *Fuzzy Risk Analysis based on Similarity Measures of Generalized Fuzzy Numbers*. Journal of IEEE Transactions on Fuzzy System 11, volume 1, 45 – 56.

Coppin, B. (2004). *Artificial Intelligence Illuminated.* 1st Edition, Jones and Bartlett Illuminated Series.

Dean, T., Allen, J., and Aloimonos, Y. (1995). *Artificial Intelligence Theory and Practice.* Addison-Wesley Publishing Company.

Falkman, G. (2000). *Similarity Measures for Structured Representations: A Definitional Approach.* Journal of Springer-Verlag Berlin Heidelberg 2000, 380-392.

Gonzalez, A. J., Xu, L., and Gupta, U. M. (1998), *Validation Techniques for Case-based Reasoning Systems.* Journal of IEEE Transactions on Systems, Man, and Cybernetics – Part A: System and Humans, volume 28, 465 – 477.

Gronlund, N. E. (2006). *Assessment of Student Achievement, 8th Edition.*

Ionescu, M., and Ralescu, A. (2005), *Evaluation of Various Aggregation Operators on Fuzzy Sets - Application to an Image Retrieval System*, Proceedings NAFIPS 2005, June 22-25, 2005, Ann Arbor, MI, USA, 670-675.

Ionescu, M., and Ralescu, A. (2005). *Fuzzy Hamming Distance Based Banknote Validator.* Proceedings of FUZZ-IEEE 2005, May 22-25, 2005 Reno, Nevada, USA, 1721-1726.

Ionescu, M., and Ralescu, A. (2004), *Fuzzy Hamming Distance in a Content-based Image Retrieval System*, Proceedings of FUZZ-IEEE 2004, Budapest, July 29-August 3. Page numbers not available

Kinley, A. (2005). *Acquiring Similarity Cases for Classification Problems*. Journal of Springer-Verlag Berlin Heidelberg 2005, 327-338.

Li, Y., Shiu, C. K., Pal, S. K., and Liu, N. K. (2005). *Learning Similarity Measure of Nominal Features in CBR Classifiers.* Journal of Springer-Verlag Berlin Heidelberg 2005, 780-785.

Martuza, V. R. (1977). *Applying Norm-Referenced and Criterion-Referenced Measurement in Education.* Allyn and Bacon, Inc., 470 Atlantic Avenue, Boston, Massachusetts 02210. Edition not has been stated.

Morell, C., Bello, R., Grau, R., and Rodriguez, Y. (2006). *Learning Similarity Metrics from Case Solution Similarity*. Journal of Springer-Verlag Berlin Heidelberg 2006, 400-408.

Negnevitsky, M. (2005). *Artificial Intelligence: A Guide to Intelligent Systems.* 2nd Edition, Pearson Education.

Nitko, A. J., and Brookhart, S. M. (2007). *Educational Assesment of Students. A Teacher Prep Text.* 5th Edition Pearson Merril Prentice Hall.

Oosterhof, A. (1994). *Classroom Applications of Educational Measurement.* 2nd Edition, Pearson Education, Inc.

Pandya, A. and Bhattacharyya, P (2005). *Text Similarity Measurement Using Concept Representation of Texts.* Pattern Recognition and Machine Intelligence. First International Conference, PReMI 2005 Kolkata, India, December 2005 Proceedings.

Popham, W. J. (2007). *Classroom Assessment: What Teachers Need to Know.* 5th Edition, Pearson Education, Inc.

Russell, S., J., and Norvig, P. (1995). *Artificial Intelligence A Modern Approach.* 1st Edition, Prentice Hall Series in Artificial Intelligence.

Stahl, A. (2005). *Learning Similarity Measures: A Formal View Based on a Generalized CBR Model.* Journal of Springer-Verlag Berlin Heidelberg 2005, German Research Center for Artificial Intelligence DFKI GmbH, Research Group Image Understanding and Pattern Recognition (IUPR), 507-521.

Teknomo, Kardi. Similarity Measurement. Retrieved 11 January, 2008 from: http://people.revoledu.com/kardi/tutorial/Similarity/index.html

Yong, D., Wenkang, S., Feng, D., and Qi, L. (2004). *A New Similarity of Generalized Fuzzy Numbers and its Application to Pattern Recognition.* Journal of Elsevire, Pattern Recognition Letters 25, 875 - 883.